Appendix C. Source and Accuracy of Estimates

SOURCE OF DATA

The data for this report were collected during the second wave of the 1987 panel of the Survey of Income and Program Participation (SIPP). The SIPP universe is the noninstitutionalized resident population of persons living in the United States. The noninstitutionalized resident population includes persons living in group quarters, such as dormitories, rooming houses, and religious group dwellings. Crew members of merchant vessels, Armed Forces personnel living in military barracks, and institutionalized persons, such as correctional facility inmates and nursing home residents, were not eligible to be in the survey. Also, United States citizens residing abroad were not eligible to be in the survey. Foreign visitors who work or attend school in this country and their families were eligible; all others were not eligible to be in the survey. With these qualifications, persons who were at least 15 years of age at the time of interview were eligible to be interviewed.

The 1987 panel SIPP sample is located in 230 primary sampling units (PSU's) each consisting of a county or a group of contiguous counties. Within these PSU's, expected clusters of two living quarters (LQ's) were systematically selected from lists of addresses prepared for the 1980 decennial census to form the bulk of the sample. To account for LQ's built within each of the sample areas after the 1980 census, a sample was drawn of permits issued for construction of residential LQ's up until shortly before the beginning of the panel. In jurisdictions that do not issue building permits, small land areas were sampled and the LQ's within were listed by field personnel, and then clusters of four LQ's were subsampled. In addition, sample LQ's were selected from supplemental frames that included LQ's identified as missed in the 1980 census and persons residing in group quarters at the time of the census.

Approximately 16,700 living quarters were designated for the sample. For Wave 1, interviews were obtained from the occupants of about 11,700 of the designated living quarters. Most of the remaining 5,000 living quarters were found to be vacant, demolished, converted to nonresidential use, or otherwise ineligible for the survey. However, approximately 800 of the 5,000 living quarters were not interviewed because the occupants refused to be interviewed, could not be found at home, were temporarily absent, or were otherwise unavailable. Thus, occupants of about 93 percent of all eligible living quarters participated in Wave 1 of the survey.

For Waves 2-7, only original sample persons (those in Wave 1 sample households and interviewed in Wave 1) and persons living with them were eligible to be interviewed. With certain restrictions, original sample persons were followed if they moved to a new address. When original sample persons moved without leaving forwarding addresses or moved to extremely remote parts of the country, additional noninterviews resulted.

Noninterviews. Tabulations in this report were drawn from interviews conducted from June through September 1987. Table C-1 summarizes information on nonresponse for the interview months in which the data used to produce this report were collected.

Table C-1. **Sample Size by Month and Interview Status**

Month	Eligible	Interviewed	Noninterviewed	Nonresponse rate (%)*
June 1987	3200	2800	400	14
July 1987	3300	2800	500	15
Aug1987	3100	2700	400	13
Sept 1987	3300	2800	500	15

*Due to rounding of all numbers at 100, there are some inconsistencies. The percentage was calculated using unrounded numbers.

Estimation. The estimation procedure used to derive SIPP person weights involved several stages of weight adjustments. Each person received a base weight equal to the inverse of his/her probability of selection. A noninterview adjustment factor was applied to the weight of every occupant of interviewed households to account for households which were eligible for the sample but were not interviewed. (Individual nonresponse within partially interviewed households was treated with imputation. No special adjustment was made for noninterviews in group quarters.) A factor was applied to each interviewed person's weight to account for the SIPP sample areas not having the same population distribution as the strata from which they were selected.

An additional stage of adjustment to persons' weights was performed to reduce the mean square error of the survey estimates by ratio adjusting SIPP sample estimates to monthly Current Population Survey (CPS) estimates[1] of the civilian (and some military) noninstitutional population of the United States by age, race, Spanish origin and sex type of householder (married, single with relatives, single without relatives, and relationship to householder (spouse or other). The CPS estimates were themselves brought into agreement with estimates from the 1980 decennial census which were adjusted to reflect births, deaths, immigration, emigration, and changes in the Armed Forces since 1980. Also, an adjustment was made so that a husband and wife within the same household were assigned equal weights.

ACCURACY OF ESTIMATES

SIPP estimates in this report are based on a sample; they may differ somewhat from the figures that would have been obtained if a complete census had been taken using the same questionnaire, instructions and enumerator. There are two types of errors possible in an estimate based on a sample survey: nonsampling and sampling. The magnitude of SIPP sampling error can be estimated, but this is not true of nonsampling error. Found below are descriptions of sources of SIPP nonsampling error, followed by a discussion of sampling error, its estimation, and its use in data analysis.

Nonsampling variability. Nonsampling errors can be attributed to many sources, e.g., inability to obtain information about all cases in the sample, definitional difficulties, differences in the interpretation of questions, inability or unwillingness on the part of the respondents to provide correct information, inability to recall information, errors made in collection such as in recording or coding the data, errors made in processing the data, errors made in estimating values for missing data, biases resulting from the differing recall periods caused by the rotation pattern and failure to represent all units within the universe (undercoverage). Quality control and edit procedures were used to reduce errors made by respondents, coders and interviewers.

Undercoverage in SIPP results from missed living quarters and missed persons within sample households. It is known that undercoverage varies with age, race, and sex. Generally, undercoverage is larger for males than for females and larger for Blacks than for non-Bblacks. Ratio estimation to independent age-race-sex population controls partially corrects for the bias due to survey undercoverage. However, biases exist in the estimates to the extent that persons in missed households or missed persons in interviewed households have different characteristics than the interviewed persons in the same age-race-Hispanic-sex group. Further, the independent population controls used have not been adjusted for undercoverage.

Some respondents do not respond to some of the questions. Therefore, the overall nonresponse rate for some items such as income and other money related items is higher than the nonresponse rates presented in table C-1. The Bureau uses complex techniques to adjust the weights for nonresponse, but the success of these techniques in avoiding bias is unknown.

Comparability with other statistics. Caution should be exercised when comparing data from this report with data from earlier SIPP products or with data from other surveys. The comparability problems are caused by sources such as the seasonal patterns for many characteristics, definitional differences, and different nonsampling errors.

Sampling variability. Standard errors indicate the magnitude of the sampling variability. They also partially measure the effect of some nonsampling errors in response and enumeration, but do not measure any systematic biases in the data. The standard errors for the most part measure the variations that occurred by chance because a sample rather than the entire population was surveyed.

Confidence intervals. The sample estimate and its standard error enable one to construct confidence intervals, ranges that would include the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these being surveyed under essentially the same conditions and using the same sample design, and if an estimate and its standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

The average estimate derived from all possible samples is or is not contained in any particular computed interval. However, for a particular sample, one can say with a specified confidence that the average estimate derived from all possible samples is included in the confidence interval.

[1]These special CPS estimates are slightly different from the published monthly CPS estimates. The differences arise from forcing counts of husbands to agree with counts of wives.

Hypothesis testing. Standard errors may also be used for hypothesis testing, a procedure for distinguishing between population parameters using sample estimates. The most common types of hypotheses tested are 1) the population parameters are identical versus 2) they are different. Tests may be performed at various levels of significance, where a level of significance is the probability of concluding that the parameters are different when, in fact, they are identical.

All statements of comparison in the report have passed a hypothesis test at the 0.10 level of significance or better. This means that, for differences cited in the report, the estimated absolute difference between parameters is greater than 1.6 times the standard error of the difference.

To perform the most common hypothesis test, compute the difference X_A - X_B, where X_A and X_B are sample estimates of the parameters of interest. A later section explains how to derive an estimate of the standard error of the difference X_A - X_B. Let that standard error b*e s_{DIFF}. If X_A - X_B is between -1.6 times s_{DIFF} and +1.6 times s_{DIFF}, no conclusion about the parameters is justified at the 10 percent significance level. If on the other hand, X_A - X_B is smaller than -1.6 times s_{DIFF} or larger than +1.6 times s_{DIFF}, the observed difference is significant at the 10 percent level. In this event, it is commonly accepted practice to say that the parameters are different. Of course, sometimes this conclusion will be wrong. When the parameters are, in fact, the same, there is a 10 percent chance of concluding that they are different.

Note concerning small estimates and small differences. Summary measures are shown in the report in tables C-2 and C-4 only when the base is 200,000 or greater. Because of the large standard errors involved, there is little chance that estimates will reveal useful information when computed on a base smaller than 200,000. Also, nonsampling error in one or more of the small number of cases providing the estimates can cause large relative error in that particular estimate. Estimated numbers are shown, however, even though the relative standard errors of these numbers are larger than those for the corresponding percentages. These smaller estimates are provided primarily to permit such combinations of the categories as serve each user's needs. Therefore, care must be taken in the interpretation of small differences since even a small amount of nonsampling error can cause a borderline difference to appear significant or not, thus distorting a seemingly valid hypothesis test.

Standard error parameters and tables and their use. Most SIPP estimates have greater standard errors than those obtained through a simple random sample of the same size because clusters of living quarters are sampled. To derive standard errors that would be applicable

Table C-2. **Standard Errors of Estimated Numbers of Persons**

(Numbers in thousands)

Size of estimate	Standard error	Size of estimate	Standard error
200	44	22,000	429
300	53	26,000	460
600	76	30,000	487
1,000	97	50,000	585
2,000	137	80,000	646
5,000	215	100,000	643
8,000	270	130,000	572
11,000	314	135,000	551
13,000	339	150,000	464
15,000	362	160,000	378
17,000	383	176,000	74

to a wide variety of estimates and could be prepared at a moderate cost, a number of approximations were required. Estimates with similar standard error behavior were grouped together and two parameters (denoted "a" and "b") were developed to approximate the standard error behavior of each group of estimates. These "a" and "b" parameters are used in estimating standard errors and vary by type of estimate and by subgroup to which the estimate applies. Table B-4 provides base "a" and "b" parameters to be used for estimates obtained from topical module data on educational attainment.

The "a" and "b" parameters may be used to calculate the standard error for estimated numbers and percentages. Because the actual standard error behavior was not identical for all estimates within a group, the standard errors computed from these parameters provide an indication of the order of magnitude of the standard error for any specific estimate. Methods for using these parameters for computation at approximate standard errors are given in the following sections. For those users who wish further simplification, we have also provided general standard errors in tables C-2 and C-3. The standard errors resulting from this simplified approach are less accurate. Methods for using these parameters and tables for computation of standard errors are given in the following sections.

Standard errors of estimated numbers. The approximate standard error, s_x, of an estimated number of persons, can be obtained in two ways. It may be obtained by use of the formula

$$s_x = s \quad (1)$$

where s is the standard error on the estimate obtained by interpolation from table C-2. Alternatively, s_x may be approximated by the formula

$$s_x = \sqrt{ax^2 + bx} \quad \mathbf{(2)}$$

from which the standard errors in table C-2 were calculated. Here x is the size of the estimate and "a"

Table C-3. **Standard Errors of Estimated Percentages of Persons**

Base of estimated percentage (thousands)	Estimated percentage					
	1 or 99	2 or 98	5 or 95	10 or 90	25 or 75	50
200	2.2	3.1	4.8	6.6	9.5	10.9
300	1.8	2.5	3.9	5.3	7.7	8.9
600	1.3	1.8	2.7	3.8	5.5	6.3
1,000	1.0	1.4	2.1	2.9	4.2	4.9
2,000	0.7	1.0	1.5	2.1	3.0	3.5
5,000	0.4	0.6	1.0	1.3	1.9	2.2
8,000	0.3	0.5	0.8	1.0	1.5	1.7
11,000	0.3	0.4	0.6	0.9	1.3	1.5
13,000	0.3	0.4	0.6	0.8	1.2	1.4
17,000	0.2	0.3	0.5	0.7	1.0	1.2
22,000	0.2	0.3	0.5	0.6	0.9	1.0
26,000	0.2	0.3	0.4	0.6	0.8	1.0
30,000	0.2	0.2	0.4	0.5	0.8	0.9
50,000	0.1	0.2	0.3	0.4	0.6	0.7
80,000	0.1	0.2	0.2	0.3	0.5	0.5
100,000	0.1	0.1	0.2	0.3	0.4	0.5
130,000	0.1	0.1	0.2	0.3	0.4	0.4
150,000	0.1	0.1	0.2	0.2	0.3	0.4
160,000	0.1	0.1	0.2	0.2	0.3	0.4
176,000	0.1	0.1	0.2	0.2	0.3	0.4

and "b" are the parameters associated with the particular type of characteristic being estimated. Use of formula 2 will provide more accurate results than the use of formula 1.

Illustration. SIPP estimates given in table 1 of the report show that there were 1,912,000 persons age 18-24 that earned a bachelor's as their highest degree. The appropriate parameters from table C-4 and the appropriate general standard error from table C-2 are

$$a = -0.000054, b = 9{,}535, s = 133{,}000$$

Using formula 1, the approximate standard error is

$$s_x = 133{,}000$$

Using formula 2, the approximate standard error is

$$\sqrt{(-0.000054)(1{,}912{,}000)^2 + (9{,}535)(1{,}912{,}000)} = 134{,}000$$

Using the standard error calculated from formula 2, the approximate 90-percent confidence interval as shown by the data is from 1,698,000 to 2,126,000. Therefore, a conclusion that the average estimate derived from all possible samples lies within a range computed in this way would be correct for roughly 90 percent of all samples.

Standard error of a mean. A mean is defined here to be the average quantity of some item. Standard errors are provided in the detailed tables for all displayed means.

Standard errors of estimated percentages. The reliability of an estimated percentage, computed using sample data for both numerator and denominator, depends upon both the size of the percentage and the size of the total upon which the percentage is based. Estimated percentages are relatively more reliable than the corresponding estimates of the numerators of the percentages, particularly if the percentages are 50 percent or more, e.g., the percent of people employed is more reliable than the estimated number of people employed. If proportions are presented instead of percentages, note that the standard error of a proportion is equal to the standard error of the corresponding percentage divided by 100. The type of percentages presented in this report is the percentage of persons sharing a particular characteristic such as the percent of persons holding a bachelor's degree.

For the percentage of persons, families, or households the approximate standard error, $s_{(x,p)}$, of the estimated percentage p can be obtained by the formula

$$s_{(x,p)} = s \qquad (3)$$

In this formula, s is the standard error on the estimate from table C-3. Alternatively, it may be approximated by the formula

$$s_{(x,p)} = \sqrt{\frac{b}{x}(p)\,(100-p} \qquad (4)$$

from which the standard errors in table 3 were calculated. Here x is the size of the subclass of social units which is the base of the percentage, while p is the percentage ($0 < p < 100$), and b is the parameter associated with the characteristic in the numerator. Use of this formula will give more accurate results than use of formula 3 above.

Table C-4. **SIPP Generalized Variance Parameters**

Characteristic	a	b
Educational attainment of all persons	-0.000054	9,535

Illustration. Text table A shows that 7.3 percent of persons age 18-24 earned a bachelor's as their highest degree. The base of this percentage is 26,148,000. Using formula 3 with the appropriate standard error from table C-3, the approximate standard error is

$$s_{(x,p)} = 0.5\,\textit{percent}$$

Using formula 4 with the "b" parameter from table 4, the approximate standard error is

$$s_{(x,p)} = \sqrt{\frac{9{,}535}{26{,}148{,}000}\,7.3\%(100\%\text{–}7.3\%)} = 0.5\%$$

Consequently, the approximate 90-percent confidence interval as shown by these data is from 6.5 to 8.1 percent.

Standard error of a difference. The standard error of a difference between two sample estimates is approximately equal to

$$s_{(x-y)} = \sqrt{s_x^2 + s_y^2} \qquad (5)$$

where s_x and s_y are the standard errors of the estimates x and y. The estimates can be numbers, percents, ratios, etc. The above formula assumes that the correlation coefficient between the two characteristics estimated by x and y is zero. If the correlation is really positive (negative), then this assumption will lead to overestimates (underestimates) of the true standard error.

Illustration. Again using text table A, 17.8 percent of persons age 25-34 earned a bachelor's as their highest degree and 15.6 percent of persons age 35-44 earned the same degree status. The bases of the percentages for persons age 25-34 and age 35-44 are 42,858,000 and 34,352,000, respectively. The standard errors for these percentages are computed using formula 4, to be .6 percent and .6 percent. Assuming that these two estimates are not correlated, the standard error of the estimated difference of 2.2 percentage points is

$$s_{(x-y)} = \sqrt{(0.6\%)^2 + (0.6\%)^2} = 0.8\%$$

Suppose that it is desired to test at the 10 percent significance level whether the percentage of persons with a bachelors as their highest degree was different for persons age 25-34 years than for persons age 35-44 years. To perform the test, compare the difference of 2.2 to the product 1.6 x .8 = 1.28. Since the difference is greater than 1.6 times the standard error of the difference, the data show that the two age groups are significantly different at the 10 percent significance level.